

Aberystwyth University

Navigational strategies and surveillance

Hogg, David C.; Dee, Hannah

Publication date:
2006

Citation for published version (APA):

Hogg, D. C., & Dee, H. (2006). *Navigational strategies and surveillance*. 73-81. <http://hdl.handle.net/2160/5876>

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

Navigational strategies and surveillance

Hannah M. Dee
Digital Imaging Research Centre
Kingston University
Kingston-Upon-Thames KT1 2EE, UK.
hannah.dee@kingston.ac.uk

David C. Hogg
School of Computing,
University of Leeds,
Leeds LS2 9JT, UK.
dch@comp.leeds.ac.uk

Abstract

This paper will describe new work that attempts to perform the modelling of human behaviour not at the level of visible patterns of motion, but at the level of intentions. By inferring intentions in terms of known goals, it becomes possible to explain the behaviour of people moving around within the field of view of a video camera (e.g. "Agent 25 went to exit 8 via sub-goals 34 and 21"). Earlier work used an ad-hoc model of human navigation and recalculated possible intentions at each frame, whereas the work presented here incorporates models of navigation from within psychology which are both simpler and more conceptually plausible, whilst providing comparable results. The basic algorithm involves generating all possible plausible paths through the scene to known goal sites, and then measuring the distance between each path and the agent's actual trajectory. Two navigational strategies are discussed, and a number of distance measures are proposed and evaluated. A prototype system has been tested on video from an outdoor car-park and an indoor foyer scene, and it has been found to produce psychologically plausible explanations in the majority of cases.

1. Introduction

Vision is a process that happens on many levels. It is possible to talk of human vision in terms of motion detection and the perception of patches of colour, but it is equally possible to talk in terms of higher level processes such as object recognition and face detection. Recent psychological studies using pre-linguistic infants have shown that we also perceive intentionality: behaviour which is clearly goal-directed is perceived differently to behaviour that isn't explicable in terms of known goals[3]. Within computer vision, the task of interpreting human behaviour from video data is a major challenge. Systems based upon machine learning techniques have been developed that learn patterns of activity over time, and hand-crafted techniques designed to spot particular behaviours have also been developed, but very little has been produced that works at the higher level

of intentionality.

1.1. Previous Work

In behaviour modelling for computer vision, three broad families of approach can be identified: learning patterns of activity over a scene [12, 16]; modelling the interactions between agents [13] or hand-crafting models of the particular behaviours of interest (e.g. [10]). A number of recent review papers[1, 9, 17] have provided strong introductions to the field of behaviour modelling and we refer the reader to these for a more thorough overview.

The crafting of detailed scene models provides an effective means of predicting behaviour and detecting atypical behaviour in a static scene, but is an approach that breaks down in situations where the possible routes through the scene are affected by changes over time (such as a car-park, or a station concourse). In such scenes the movement of obstacles can have a great effect upon the paths people take. Approaches which concentrate on the interactions between agents suffer from a related problem: sometimes the agents interact with the scene rather than each other. Modelling specific behaviour patterns is a costly way to create a system, and whilst useful for dealing with specific alarm situations (graffiti or abandoned package detection, for example) is of less use when modelling wider ranges of human behaviour.

The approach we propose is inspired by the work of Dennett[6], and involves taking a step back from the physical environment and instead modelling the psychology of the agents within the scene. By considering the behaviour of the people moving around the scene as the behaviour of *intentional agents*, we can reason about their behaviour in terms of goals and sub-goals. This leads to a number of advantages over existing systems: the easy summarisation of behaviour (*Person N is going towards goal M*); the ability to deal with a changing layout of goals (such as one finds in a car-park); the ability to deal with uncommon but not unrealistic paths (such as rarely used short-cuts); and the ability to deal with flexible scenes without well-defined paths. The approach we propose involves constructing a simple model

of scene (exits and obstacles), and then working out which of the exits (goals) within the scene are plausible for a particular agent according to some model of human navigation.

1.2 Human navigational strategies

As this work attempts to model human intentional behaviour in environments containing obstacles, it is important that we take into consideration the way in which people actually navigate around a scene. The question addressed in this section is whether or not our day-to-day path planning activity is rational (do we actually take the shortest or least-cost path?) and by investigating this question at the outset, we hope our subsequent models are able to reflect more accurately the way in which people actually plan their route. Naïvely, people could be expected to take the shortest route from A to B, however various other factors also seem to influence navigational decisions.

Golledge [7] found that asymmetries in planning (i.e., coming back via a different route) were common: indeed, in one of his scenarios, 75% of subjects returned via a different route. In [8] Golledge ranks the different strategies used in path planning (shown in box). The fact that Golledge found “fewest turns” or *simplest path* to be one of the most attractive metrics for path planning implies that either our distance perception is skewed by path complication, or there is some other reason for preferring simplest paths (for example, some form of cost is associated with changes in direction). It has also been suggested that distances are perceived as being shorter or longer dependant upon whether they proceed away from or towards primary route nodes or reference points (Sadalla et al [14]). This is borne out by anecdotal accounts of people taking different routes from and to a particular place. Conroy-Dalton [2] calls this the British Library effect, after the place in which she observed the behaviour. This could be due, as Conroy-Dalton suggests, to a preference for the straightest path between A and B, or it could be due to wanting to make the first stretch of a journey the most significant (in terms of distance travelled towards final goal).

2 Projecting ideal paths

To facilitate the construction of potential paths through the scene we construct an agent-centered “map” or representation of the scene. This starts with the determination of the area of scene directly visible¹ to the agent. In a scene without obstacles, the representation would stop here: it is

¹The models described here assume that the agents *know where they are going* and are not engaged in a random walk: this seems to be a fair assumption, as even in totally unknown situations, people are more likely to have consulted a map or have asked directions than to search with no knowledge of their final goal location.

Ranking of path planning strategies, from [8]

1. Shortest distance
2. Least time
3. Fewest turns
4. Most scenic/aesthetic
5. First noticed
6. Longest leg first
7. Many curves
8. Many turns
9. Different from previous (novelty)
10. Shortest leg first

possible to determine which goals are consistent with the movement of an agent by working out which they are moving towards. With obstacles, the problem becomes more interesting. The agents’ actual goal may be obscured, requiring the agent first to move away from the goal in question in order to circumnavigate some intervening obstacle. In order to account for this behaviour, we use virtual “*sub-goals*”, which are defined as points in the scene where an agent might choose to change direction.

The construction of sub-goals is based upon geographical information about the location of obstacles, the current location of the agent within the scene x and their direction of motion θ , and upon counterfactual reasoning. With obstacle location recorded in a polygonal form (this map is hand-crafted) we consider each tangential obstacle vertex in turn². If it is visible from x (that is, if the line from x to v does not pass through any other obstacles) it is a sub-goal. Each of these directly visible sub-goals is then investigated in turn, looking for further sub-goals visible from there.

Using these agent centred maps as a starting point, all plausible paths to known goals are generated according to some model of navigational strategy. This results in a tree of possible future routes through the scene, all of which are consistent and “reasonable” for a rational agent. What remains to be done is to determine which of these possible ideal paths the agent is most likely to be following, and then to determine some degree of fit. If the degree of fit is poor, then that path is not a good explanation for that agent’s behaviour. The two models of human navigation we investigate in the following sections are *Simplest Path*, in which routes are planned using as few sub-goals as possible, and *Shortest Path*, in which routes are planned minimizing total proposed travel distance. All known exits to the scenes are considered, and if there exists a path from the agent to that exit, then that path is added to the tree of possible paths. In many cases, *shortest path* and *simplest path* result in the same path – especially in simple scenes. Shortest and sim-

²A tangential vertex is one whose neighbouring vertices are on the same side of a line through the agent’s location and the vertex in question

plest paths for an example case are shown in Figures 1 and 2.

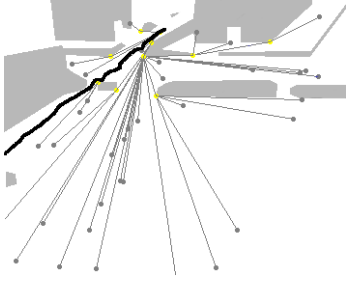


Figure 1: All shortest paths for a sample agent: Actual trajectory in black, hypothetical paths in dark grey, obstacles in light grey.

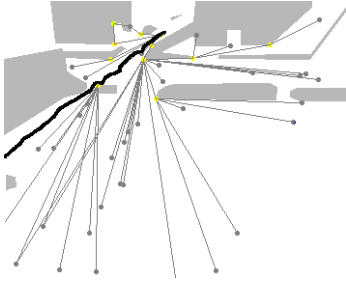


Figure 2: All simplest paths for a sample agent: colour code as in Figure 1.

3 Trajectory distance functions

Given these trees of possible paths (two for each agent - one tree representing the *Simplest paths* to each known goal, and one tree representing the *Shortest paths*), the next task is to determine which of the paths the agent is following, and how closely they are following that path. Each path can be thought of as a possible intention – at this instance in time, *Agent A might be headed towards goal X via sub-goals n and m*. The job is to work out which of these possible intentions fits the visible behaviour of the agent.

The Hausdorff distance h is a measure of distance from a set of points X to a second set of points Y and is defined as the maximum distance of a set of points to the nearest point in the other set.

$$h(X, Y) = \max_{x \in X} \{ \min_{y \in Y} \{ \|x - y\| \} \} \quad (1)$$

This is an asymmetric measurement, and it is common for authors to take the Hausdorff distance *between* two sets

- that is, to calculate the distance from each set to the other, and take the higher of the two. There are two problems with using the Hausdorff measure between two sets in the current case. Firstly, the hypothetical paths are represented by a small number of points (a path with n sub-goals will be defined by $n + 2$ points: a point at the start and end points, and one for each sub-goal) whilst the trajectories have a set number of points for each second the agent was within the field of view of the camera. This problem is avoided by first quantizing each ideal path so that it is defined using m points, where m is the number of frames in the corresponding portion of trajectory. Secondly, the Hausdorff distance measures the distance between two sets of points, whilst the trajectories and paths under examination are point sequences. The correspondences considered by the Hausdorff distance will not in general respect this ordering, as is shown in Figure 3.

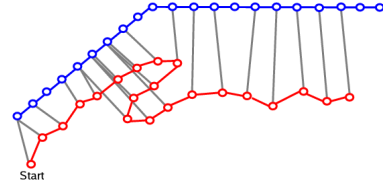


Figure 3: The selection of matched points in Hausdorff calculations.

We propose an alteration of the Hausdorff distance to impose sequential ordering upon the permitted point correspondences. The proposed *monotonic* Hausdorff distance is a distance from one sequence to another. Here, the minimisation is over monotonically increasing functions $\mathcal{I} : \{1, 2 \dots m\} \rightarrow \{1, 2 \dots m\}$, thereby imposing the required constraint.

$$h_m \{ \{x_1, x_2 \dots x_n\}, \{y_1, y_2 \dots y_n\} \} = \max_{I \in \mathcal{I}} \left(\min_i \|y_{I(i)} - x_i\| \right) \quad (2)$$

A diagram illustrating the difference in the selection of which point to match is shown in Figure 4. It is clear from Figure 3 that with the standard Hausdorff distance, low matches can occur in situations where the agent doubles back upon themselves (as there is always a point on the idealised path near to the trajectory). Figure 4 shows that by forcing the matched point to have a monotonically increasing distance from the start of the ideal path, this problem is avoided.

The monotonic Hausdorff distance measure is inadequate by itself as a measure of intentionality: whilst it identifies which path the agent is likely to be following by determining the closest path in space, it does not provide an ideal measure of goal-directedness. Given that the agent is

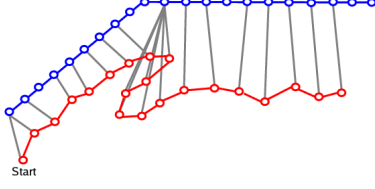


Figure 4: The selection of matched points in Monotonic Hausdorff calculations.

following the closest path according to the monotonic Hausdorff distance a better measure of goal directedness is obtained from measures that capture angular disparity from the shortest or simplest path (penalising meandering paths). Hence we use a two stage process as shown in Figure 5. There are three proposed distance measures for the final measure of goal-directedness: AD , IS and C . All of these measures rely to some extent on angular disparity.

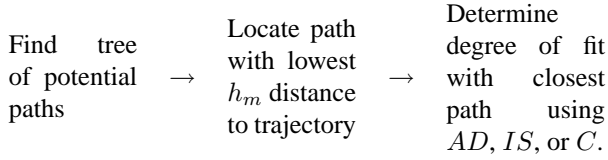


Figure 5: Overview of algorithm.

Each hypothesised path (either simplest or shortest) is constructed as a number of straight line segments. Taking the angle of each of these segments ϕ and comparing it to the angle of motion of the agent θ provides us with an indication of whether the agent is moving in the general direction of the goal or sub-goal. Figures 6 and 7 show the trajectories of the agent (in black), and the monotonic Hausdorff distance closest shortest path (in grey). Also shown is a graph of the angular disparity between the agents' trajectory at each time step and each of the segments of the projected best path.

The agent whose trajectory and ideal path are shown in Figure 6 travels very smoothly through the scene. The graph of angular disparity in this case is very clear: at any one point it is obvious which section has the lowest angular disparity (although it is worth noting that the noise at the start of the trajectory before the Kalman filter has stabilised is clear to see). The agent depicted in 7 shows a pedestrian. Two of the three sections of the idealised path have clearly different orientations, and choosing the point where the agent transitions from one to the next is straightforward based upon angle alone. The trajectory, however, is noisy (due to the up-and-down bobbing motion of a walking person) yet clearly goal-directed.

By examining the angular disparity between the velocity

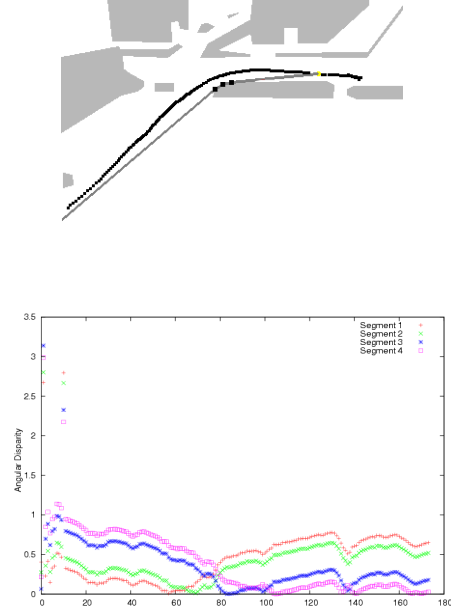


Figure 6: Car trajectory identified as having 4 segments, pictured above its corresponding angular disparity graph. The sub-goals are indicated on the trajectory image with black dots. The angular disparity graph shows the difference in angle between the direction of the agent and each of the four segments throughout the length of the agent's trajectory. From this graph, it is clear to see that the agent's trajectory is closest in direction to segment 1 at the start of the trajectory, then closest to segment 2, then 3 and finally 4. This is as expected, as the predicted closest path goes through these segments in that order.

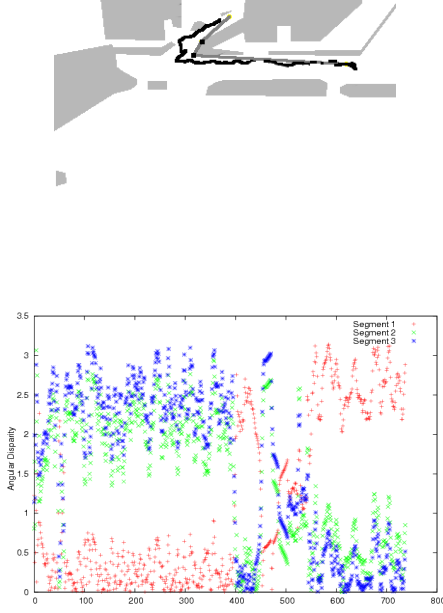


Figure 7: Pedestrian trajectory identified as having 3 segments. The graph shown indicates that for the first part of the trajectory, the agent is heading in the same direction as segment 1. Segments 2 and 3 are harder to separate as the trajectory is fairly noisy.

of the agent and each of the segments of the ideal path identified as being closest, it is possible to determine at which point the agent moves from following one segment to following the next. These transition points between segments represent changes in the currently active goal of the agent: they move from heading towards a sub-goal to heading towards a goal, (or from one sub-goal to the next sub-goal). For example, the trajectory shown in Figure 6 is made up of four linear segments, and the graph shown in Figure 6 supports this. It is possible to use angular disparity to work out where the transition in goals falls – at which point the agent reaches a sub-goal.

The trajectory could be partitioned (and “actual” sub-goal location found) by minimising over all possible sets of segment transition times the modulus of the total angular disparity between the direction θ_i of the agent at that time step and the direction of the corresponding segment ϕ_k as shown in Equation 3³. This angular distance function (hereafter AD) serves two purposes: the act of minimizing partitions the trajectory by placing virtual turning points at times (v_k) that correspond to changes in direction (sub-goals), and the final value provides a measure of fit between the trajectory and the path. \mathcal{V} is the set of all sequences of n transition

³It is worth noting that the angular disparity is not, strictly speaking, a subtraction: it is the acute angular disparity between the two directions and shown here in the equations as a subtraction for simplicity.

times with $v_i = 1$ and $v_{k+1} = m + 1$ (m is the number of frames). So in a trajectory such as that shown in Figure 7, with three segments, this function has the effect of dividing the agent’s trajectory into 3 segments ($n = 3$) based upon the direction of travel of the agent and the direction of each of the predicted path segments $\theta - \phi$. In the current implementation, minimisation is carried out by performing an exhaustive search over all \mathcal{V} . This is feasible as the majority of ideal paths in the scenes used for testing have fewer than 5 segments. Within a larger search space or more complicated scenes, techniques such as Dynamic Programming would provide a faster solution.

$$AD = \min_{v \in \mathcal{V}} \sum_{k=1}^n \sum_{i=v_k}^{v_{k+1}-1} \frac{|\theta_i - \phi_k|}{m\pi} \quad (3)$$

In the majority of cases, this approach finds plausible locations for the change in direction: they fall near the sub-goals on the ideal path.

The second distance measure we consider includes a penalty term taking into account the proportion of the trajectory assigned to each segment, as well as the angular disparity term shown in Equation 3. Equation 4 shows this modified distance measure. The second component of this measure is a term which implies a penalty if the proportion of trajectory assigned to each path segment is not similar. In this, p_i represents the length of the i th path segment, and s_k is the length of the k^{th} trajectory “segment” as shown in Equation 5. The additional term of Equation 4, therefore, compares the length of corresponding segments as proportions of the total length of the path or trajectory. λ is a weighting term which is to be determined experimentally, and a value of 0.01 has been chosen for all experiments. This measure we call *Cost*, or C .

$$C = \min_{v \in \mathcal{V}} \sum_{k=1}^n \left[\sum_{i=v_k}^{v_{k+1}-1} \left(\frac{|\theta_i - \phi_k|}{m\pi} \right) + \left(\lambda \left| \frac{s_k}{\sum_n s_j} - \frac{p_k}{\sum_n p_j} \right| \right) \right] \quad (4)$$

$$\text{where } s_k = \sum_{i=v_k}^{v_{k+1}-1} \|x_i - x_{i+1}\| \quad (5)$$

The third measure we consider is the result of relaxing the criterion for goal-directedness, and considering an agent as headed towards a goal if the goal falls within half a radian either side of the agent’s velocity vector (as in Equations 7 and 6). This is computed by subtracting 0.5 radians from the angular disparity, summing over the length of the trajectory, but ignoring negative results. This final approach provides a more robust indication of goal-directedness. This measure

is called *Angular disparity Ignoring Small angles*, and we have abbreviated this to *IS*.

$$IS = \min_{v \in \mathcal{V}} \sum_{k=1}^n \sum_{i=v_k}^{v_{k+1}-1} \frac{|\tau(\theta_i - \phi_k)|}{m\pi} \quad (6)$$

$$\text{where } \tau(x) = \begin{cases} x & \text{if } x > 0.5 \\ 0 & \text{if } x \leq 0.5 \end{cases} \quad (7)$$

4 Results and evaluation

The method has been tested on two scenes: one an outdoor car-park scene featuring natural behaviours, and one an indoor foyer scene with all behaviours performed by actors. Agents in the outdoor scene were tracked using a generic blob tracker[11] with some hand-editing of tracks for consistency, and the indoor scene tracks were provided alongside the video as part of the PETS2004 workshop⁴. The exit model was learnt for the outdoor scene by fitting a Gaussian mixture model to the start and end points of trajectories, and hand-crafted for the indoor scene. Illustrations of exit models, obstacle models and scenes are shown in Figure 8. The car-park scene features nearly 300 agents over the course of an hour, and the PETS2004 scene is made up of short sequences of video featuring the same actors, which results in 24 individual trajectories. Further details on these datasets are available from[4]. The velocity component of a Kalman filter was used in order to gain an indication of each agent’s heading, and ideal paths were generated 30 frames into each agent’s trajectory (this is near the beginning, but allows time for the Kalman filter to stabilise).

As mentioned earlier, shortest and simplest paths are often the same. Of the 9,675 shortest paths predicted over the entire car-park dataset, approximately 80% (7,824) were the same as one of the corresponding set of simplest paths. There were a few cases in the car-park dataset of agents who were not heading towards any goals or sub-goals at the 30th time-step, and these agents were given scores on each of the distance measures just higher than the maximum. There is a sense in which the behaviour of these agents can be thought of as totally inexplicable within this framework: their behaviour was not goal-directed, or at least it was not goal-directed in the simple geographical sense we use here.

A visual inspection of the actual paths and the “best explanations” provided by this method confirmed that in general, the paths and trajectories had a good fit and thus the explanations were plausible. Some sample explanations are shown in Figures 9 and 10. In Figure 9 an agent from the PETS2004 dataset is shown. This is an example of an agent for whom *shortest path* and *simplest path* provide different explanations: the closest match for each is a different route

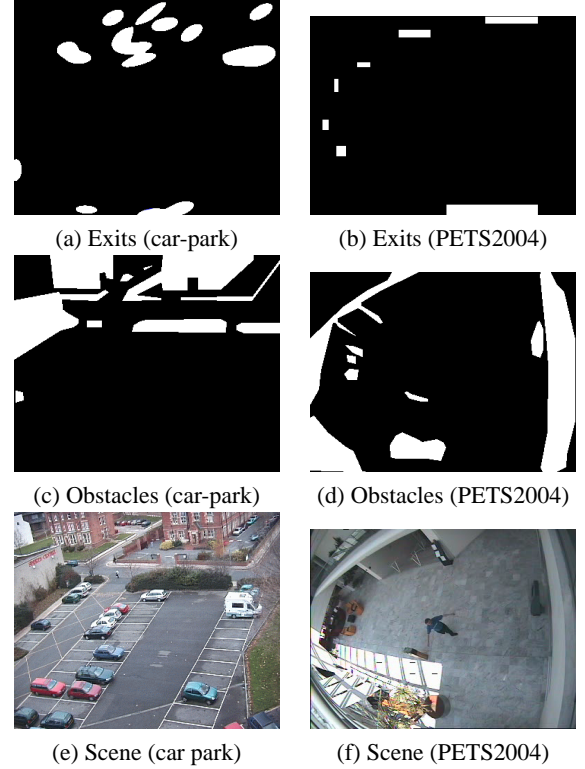


Figure 8: The exit model, obstacle model and scene for the car-park (left) and PETS2004 (right) scenes.

through the scene. This is also an example of an agent for whom the match between agent and path is fairly poor: the distance measures (*IS*, *AD* and *C*) are all fairly high (above the mean in all cases). These high scores can be interpreted as indicating that the best explanation for the behaviour is not a particularly *good* explanation – the behaviour pattern is inexplicable within our framework.

Figure 10 shows an example agent from the car-park dataset for whom *shortest path* and *simplest path* predict the same route. The various distance functions (*AD*, *IS* and *C*) are high for this agent too: the explanation provided is not very good. For examples of agents where the explanation provided *was* good, consider agents 36 and 44 from the car-park dataset, shown in Figures 7 and 6 earlier in the paper. The various distance functions for Agent 36 (where the fit between path and trajectory is quite good) are: *C* = 0.139; *AD* = 0.438; *IS* = 0.143. For Agent 44 where the explanation is a stronger fit, the distance measures are lower still: *C* = 0.025; *AD* = 0.0787; *IS* = 0.

This work has been evaluated as described in [5], in which a system for benchmarking surveillance applications against the judgements of a group of humans is described. This involves showing a group of volunteers the surveillance videos in question and asking them to rate the be-

⁴Generated as part of the EC Funded CAVIAR project/IST 2001 37540.

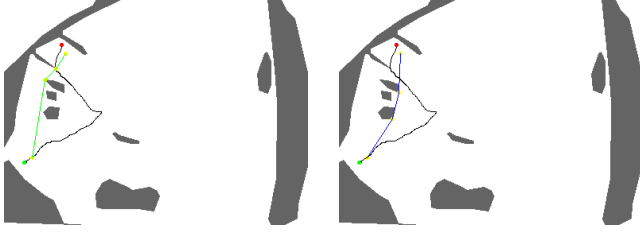


Figure 9: Agent 4 is heading towards goal 4 according to shortest path (left) C : 0.196; AD : 0.603; IS : 0.212. Simplest path predicts a different path. (right) C : 0.248; AD : 1.02; IS : 0.546



Figure 10: Agent 28 is heading towards goal 32 (shortest and simplest path) C : 0.24; AD : 0.755; IS : 0.314

haviour of each agent on a scale of 1 to 5, according to the following instructions:

“If you were a security guard, would you regard the behaviour of the agent highlighted in this video as interesting? Please indicate on the following questionnaire, with one being uninteresting and five being interesting.”

The resultant human scores can be averaged, and then the computer generated distance measures can be correlated with these human ranks in order to determine whether or not those behaviour patterns found interesting by humans are also those which score highly according to our measures. This provides a means of evaluating surveillance software and systems that avoids the problem of relying on actors or serendipity to provide “interesting” or “unusual” events. Correlations have been calculated using Kendall’s Tau T_k [15]⁵.

$$T_k = \frac{\text{concordant} - \text{discordant}}{\sqrt{n(n-1) - T_x} \sqrt{n(n-1) - T_y}} \quad (8)$$

In Equation 8, T_x and T_y are the terms correcting for tied ranks. The distribution for T_k is known for the null hypothesis of no relationship between variables. Indeed, T_k is approximately normally distributed and using Equation 9[15]

⁵Spearman’s Rho results in similar levels of significance as expected

can be converted to z scores then compared with standard statistical tables in cases where $n > 10$.

$$z = \frac{3T_k \sqrt{n(n-1)}}{\sqrt{2(2n+5)}} \quad (9)$$

For the car-park dataset, all correlations between distance functions and human ranks were positive and significant at the 0.0001 (0.01%) level. For the PETS2004 dataset, all correlations were positive and significant at the lower 0.005 level (0.5%). We do not believe that these lower levels of correlation in the PETS2004 case are problematic. The correlations are still very strong indeed, and the PETS2004 dataset is a contrived dataset made up entirely of behaviours performed by actors.

Comparing the three measures, the most highly correlated is IS : angular disparity ignoring small angles. C and the simple angular disparity function (AD) perform similarly. Shortest path correlates with the human ranks as well or better than simplest path in all cases. This supports our intuition that shortest path provided more plausible ideal paths than simplest path when looking at the explanations provided by the system. It also supports the evidence of psychologists, who have found that shortest path is the most popular route selection criteria. Figure 11 shows a chart of Kendall’s Tau (T_k) for each measure for each strategy within the car-park dataset, and Figure 12 shows the same information for the PETS2004 dataset.

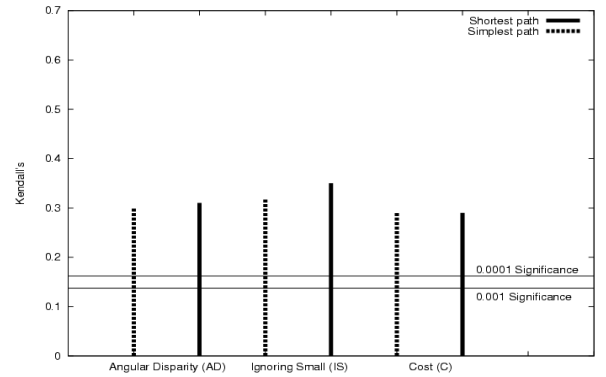


Figure 11: Summary of correlation statistics for the car-park dataset.

5 Summary and Conclusions

This paper has described a novel approach to behaviour modelling for surveillance. By abstracting away from the scene in question and concentrating instead upon modelling the psychology of the people within the scene, the system described can handle unusual short cuts and novel paths

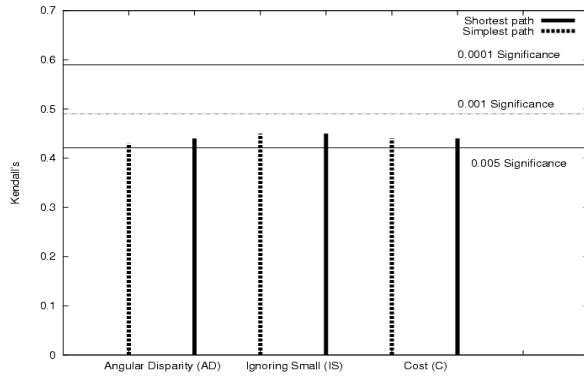


Figure 12: Summary of correlation statistics for the PETS2004 dataset.

(that is, previously unseen but goal-directed paths receive a low score on all of our metrics). The system also copes with a changing goal dynamic: as cars park, more goals appear, and as cars drive off so do the goals they represent. Moving obstacles have not yet been implemented, but the framework described here should allow for their incorporation with ease. Various distance measures have been demonstrated to correlate with human judgements of interestingness from a surveillance perspective. A further contribution of this paper is to provide a framework within which naturalistic experiments into path-planning and route selection could be conducted. The comparison between shortest path and simplest path described here confirms the strategies' relative ranking according to Golledge[8], and the investigation of other navigational strategies is left for future work.

References

- [1] Hilary Buxton. Learning and understanding dynamic scene activity: a review. *Image and Vision Computing*, 21(1):125–136, 2003.
- [2] R. Conroy Dalton. The secret is to follow your nose: Route path selection and angularity. *Environment and Behavior*, 35(1):107–131, 2003.
- [3] G. Csibra, G. Gergely, S. Bíró, O. Koós, and M. Brockbank. Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition*, 72:237–267, 1999.
- [4] H. M. Dee and D. C. Hogg. Detecting inexplicable behaviour. In *Proc. British Machine Vision Conference (BMVC)*, Kingston-on-Thames, UK, 2004.
- [5] H. M. Dee and D. C. Hogg. Is it interesting? comparing human and machine judgements on the PETS dataset. In *ECCV-PETS: the Performance Evaluation of Tracking and Surveillance workshop at the European Conference on Computer Vision*, Prague, Czech Republic, 2004.
- [6] D. C. Dennett. *The Intentional Stance*. The MIT Press/Bradford Books, Cambridge, MA, 1987, reprinted 2002.
- [7] R. Golledge. Path selection and route preference in human navigation: a progress report. In *Proc. Spatial Information Theory: Foundations of GIS (COSIT)*, pages 207–222, 1995.
- [8] R. G. Golledge. Defining the criteria used in path selection. Technical Report UCTC No. 78, University of California Transportation Center, 1995.
- [9] W. Hu, T. Tan, L. Wang, and S. Maybank. A survey on visual surveillance of object motion and behaviours. *IEEE transactions on Systems, Man and Cybernetics*, 34(3):334–352, 2004.
- [10] T. Jan, M. Piccardi, and T. Hintz. Detection of suspicious pedestrian behavior using modified probabilistic neural network. In *Proc. of Image and Vision Computing*, pages 237–241, Auckland, New Zealand, 2002.
- [11] D. R. Magee. Tracking multiple vehicles using foreground, background and shape models. *Image and Vision Computing*, 22:143–155, 2004.
- [12] D. Makris and T. Ellis. Learning semantic scene models from observing activity in visual surveillance. *IEEE transactions on Systems, Man and Cybernetics*, 35(3):397–408, 2005.
- [13] N. M. Oliver, B. Rosario, and A. P. Pentland. A Bayesian computer system for modeling human interactions. *IEEE transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 22(8):831–843, 2000.
- [14] E. K. Sadalla, W. J. Burroughs, and Staplin L. J. Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory*, 6(5):516–28, 1980.
- [15] S. Siegal and N. J. Castellan. *Nonparametric statistics for the behavioral sciences*. McGraw Hill, Singapore, 2nd edition, 1988.
- [16] Chris Stauffer and Eric Grimson. Learning patterns of activity using real-time tracking. *IEEE transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 22(8):747–757, 2000.
- [17] M. Valera and S. A. Velastin. Intelligent distributed surveillance systems: a review. *IEE. Proceedings Vision Image and Signal Processing*, 152(2):192–204, 2005.